(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: KENT, Stephen, B., H. [NZ/US]; 6046 La Jolla Hermosa Avenue, La Jolla, CA 92037 (US). MUIR, Tom, W. [GB/US]; Apartment 15B, 500 E. 63rd Street, New York, NY 10021 (US). DAWSON, Philip, E. [US/US]; 1302 Lyndon Street #11, South Pasadena, CA 91030 (US). BARK, Steven, J. [US/US]; 7944 Caminito Dia #1, San Diego, CA 92037 (US). CANNE, Lynne [US/US]; 1103 Crespi Drive, Pacifica, CA 94044 (US).

(74) Agents: LEWIS, Donald, G. et al.; The Scripps Research Institute, 10550 North Torrey Pines Road, TPC–8, La Jolla, CA 92037 (US).

(54) Title: GENERAL CHEMICAL LIGATION

(57) Abstract

Oligopeptides are ligated to form ligation peptide products. In the first step, two starting oligopeptides are ligated to form an intermediate having an aminothioester linkage. In the second step, the aminothioester linkage undergoes a rearrangement to form a peptide having an *N*–substituted amide linkage. In an optional third step, the *N*–substitution of the amide linkage is chemically removed to form a native peptide linkage.

## FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| Code | Country | Code | Country | Code | Country | Code | Country |
|------|---------|------|---------|------|---------|------|---------|
| AL | Albania | ES | Spain | LS | Lesotho | SI | Slovenia |
| AM | Armenia | FI | Finland | LT | Lithuania | SK | Slovakia |
| AT | Austria | FR | France | LU | Luxembourg | SN | Senegal |
| AU | Australia | GA | Gabon | LV | Latvia | SZ | Swaziland |
| AZ | Azerbaijan | GB | United Kingdom | MC | Monaco | TD | Chad |
| BA | Bosnia and Herzegovina | GE | Georgia | MD | Republic of Moldova | TG | Togo |
| BB | Barbados | GH | Ghana | MG | Madagascar | TJ | Tajikistan |
| BE | Belgium | GN | Guinea | MK | The former Yugoslav | TM | Turkmenistan |
| BF | Burkina Faso | GR | Greece | | Republic of Macedonia | TR | Turkey |
| BG | Bulgaria | HU | Hungary | ML | Mali | TT | Trinidad and Tobago |
| BJ | Benin | IE | Ireland | MN | Mongolia | UA | Ukraine |
| BR | Brazil | IL | Israel | MR | Mauritania | UG | Uganda |
| BY | Belarus | IS | Iceland | MW | Malawi | US | United States of America |
| CA | Canada | IT | Italy | MX | Mexico | UZ | Uzbekistan |
| CF | Central African Republic | JP | Japan | NE | Niger | VN | Viet Nam |
| CG | Congo | KE | Kenya | NL | Netherlands | YU | Yugoslavia |
| CH | Switzerland | KG | Kyrgyzstan | NO | Norway | ZW | Zimbabwe |
| CI | Côte d'Ivoire | KP | Democratic People's | NZ | New Zealand | | |
| CM | Cameroon | | Republic of Korea | PL | Poland | | |
| CN | China | KR | Republic of Korea | PT | Portugal | | |
| CU | Cuba | KZ | Kazakstan | RO | Romania | | |
| CZ | Czech Republic | LC | Saint Lucia | RU | Russian Federation | | |
| DE | Germany | LI | Liechtenstein | SD | Sudan | | |
| DK | Denmark | LK | Sri Lanka | SE | Sweden | | |
| EE | Estonia | LR | Liberia | SG | Singapore | | |

## GENERAL CHEMICAL LIGATION

### Description

### Technical Field of the Invention:

The present invention relates to methods for
chemically ligating oligopeptides.  More particularly,
the invention relates to methods for chemically ligating
oligopeptides end to end using a peptide bond.

### Government Rights:

This invention was made with government support
under Grants No. R01 GM 48897 and P01 GM 48870 awarded
by the National Institutes of Health.  The U.S.
government has certain rights in the invention.

### Background of the Invention:

The synthesis of peptides by conventional stepwise
solid phase methodologies is limited by low yields when
synthesizing long peptides. (Merrifield et al. *J. Am.
Chem. Soc.* **1963**, *85*, 2149-2154; Kent et al. *Ann. Rev.
Biochem.* **1988**, *57*, 957-989.)  To overcome this
limitation, smaller synthetic peptides may be joined to
one another by chemical ligation to produce longer
peptide products.

A method for chemically ligating peptides was
disclosed by Schnölzer et al.. (Schnölzer et al.,
*Science* **1992**, *256*, 221-225; Rose et al. *J. Am. Chem.
Soc.* **1994**, *116*, 30-34; Liu et al. *Proc. Natl. Acad. Sci.
USA* **1994**, *91*, 6584-6588).  The methodology disclosed by
Schnölzer involves a chemoselective reaction of
unprotected peptide segments to give a product with an
**unnatural** backbone structure at the site of ligation.
This methodology makes possible the synthesis of
peptides of greater size than attainable by standard
methods of peptide synthesis. (Canne et al. *J. Am. Chem.*

Soc. **1995**, *117*, 2998-3007; Baca et al. *J. Am. Chem. Soc.* **1995**, *117*, 1881-1887; Williams et al. *J. Am. Chem. Soc.* **1994**, *116*, 10797-10798). This methodology also makes possible the synthesis of peptides of unusual structure and topology. (Dawson et al. *J. Am. Chem. Soc.* **1993**, *115*, 7263-7266; Rose et al. *J. Am. Chem. Soc.* **1994**, *116*, 30-34; Muir et al. *Biochemistry* **1994**, *33*, 7701-7708; Canne et al. *J. Am. Chem. Soc.* **1995**, *117*, 2998-3007). The combined use of conventional stepwise solid phase peptide synthesis together with chemical ligation enables chemists to routinely make unprotected peptides of up to 60 amino acid residues in good yield and purity. (Schnölzer et al. *Int. J. Pept. Protein Res.* **1992**, *40*, 180-193). The combination of these two methodologies may also be employed to achieve a total chemical synthesis of proteins.

Another chemical ligation technique has been reported for the preparation of proteins having a native backbone structure (Dawson et al. *Science* **1994**, *266*, 776-779). This mode of chemical ligation is termed "native ligation." In this technique, an unprotected synthetic peptide bearing a *C*-terminal α-thioester is reacted in a chemoselective manner with an unprotected peptide containing an *N*-terminal Cys residue. Thiol exchange reaction yields an intial thioester-linked intermediate which spontaneously rearranges to give a native amide bond at the ligation site joining the two peptide segments, in the process regenerating the Cys side chain thiol. This version of native ligation uses chemistry first described by Wieland for reacting amino acids. (Wieland et al. *Liebigs Ann. Chem.* **1953**, *583*, 129-149.) As originally described, native ligation is restricted to joining peptide segments at an **X**-Cys bond. (Dawson et al. *Science* **1994**, *266*, 776-779.)

What is needed is a general method for joining a $C$-terminal $\alpha$-thioester peptide segment to an $N$-terminal amino acid peptide segment, wherein the $N$-terminal amino acid peptide segment need not have an $N$-terminal

5      cysteine.


### Summary of the Invention:

The invention is directed to a method for

10     chemically ligating unprotected oligopeptides to form a product having all peptide linkages. In the first step, two oligopeptides are ligated to form a ligation product having an aminothioester linkage. In the second step, the aminothioester linkage rearranges to form a product

15     having an $N$-substituted amide linkage. In an optional third step, the substitution on the amide bond is removed by facile treatment with Zn in acidic medium, to give a native peptide bond at the ligation site.


20     The method employs two starting oligopeptides, viz., a first starting oligopeptide and a second starting oligopeptide. The first oligopeptide has a $C$-terminal auxiliary group with a thioester moiety, i.e., $[peptide_1]^{\alpha}COSR$, where R is selected from the group

25     consisting of 3-carboxy 4-nitrophenyl and benzyl. The second oligopeptide has an $N$-terminal auxiliary functional group with an unoxidized sulfhydryl moiety, i.e., $HSCH_2CH_2(O)-N^{\alpha}[peptide_2]$. When the first and second starting oligopeptides are admixed under conditions

30     promoting thioester exchange, they condense with one another to form an intermediate oligopeptide product wherein the first and second oligopeptides are linked via an amino-thioester bond. The amino-thioester bond then spontaneously rearranges intramolecularly to form a

35     ligation product linked by an $N$-substituted amide bond. During the intramolecular rearrangement, the amino group of the $N$-terminal auxiliary functional group atacks the

thioester to form an amide bond with the attached *N*-linked auxiliary functional group containing a displaced sulfhydryl moiety.  The *N*-linked auxiliary functional group containing the displaced sulfhydryl moiety may

5      then be optionally removed by chemical means to form a product having all native peptide linkages.

One aspect of the invention is directed to a method for ligating a first oligopeptide with a second

10     oligopeptide end to end for producing an oligopeptide product.  More particularly, the method comprises two steps with an optional third step.

The first step involves condensing the *C*-terminal

15     thioester of a first peptide with the unoxidized sulfhydryl moiety of a second oligopeptide for producing an intermediate oligopeptide linking the first and second oligopeptides with an δ or γ - amino-thioester bond.  The first oligopeptide includes the *C*-terminal

20     thioester on a *C*-terminal residue and the second oligopeptide includes the *N*-terminal auxiliary functional group on an *N*-terminal residue having an unoxidized sulfhydryl moiety.  If the *C*-terminal residue on the first oligopeptide is non glycine then the *N*-

25     terminal residue on the second oligopeptide is glycine and if the *N*-terminal residue on the second oligopeptide is non glycine then the *C*-terminal residue on the first oligopeptide is glycine with the proviso that the non glycine residue is a non β-branched amino acid.

30

The second step rearranges the δ or γ - amino-thioester bond of the intermediate oligopeptide via intramolecular attack of the δ or γ - amino group onto the thioester moiety and displaces a sulfhydryl moiety

35     as a byproduct from the thioester moiety thereby producing an oligopeptide product linking the first and second oligopeptides with an amide bond.  The nitrogen

of the amide bond contains the auxiliary functional
group with the displaced sulfhydryl moiety.

An optional third step involves the removal of the
auxiliary functional group on the amide nitrogen, from
the oligopeptide product with a reducing agent for
producing a native peptide bond. The amide nitrogen
auxiliary functional group is $N$-$\alpha$-O-$(CH_2)_n$-SH wherein
$1 \le n \le 2$ and the reducing agent is Zinc (underlined
nitrogen represents the coupled amide nitrogen).

Another apect of the invention includes the
oligopeptide intermediate comprising a first
oligopeptide segment including a $C$-terminal thioester, a
second oligopeptide segment including an $N$-terminal
auxiliary functional group having an unoxidized
sulfhydryl moiety, and an aminothioester linkage unit
linking the $C$-terminal thioester and the sulfhydryl
group of the auxiliary functional group.

## Description of Figures:

Figure 1 illustrates the generalized native
chemical ligation of unprotected peptide segments
wherein **R** is a non $\beta$-branched amino acid residue with
the following conditions: if the $C$-terminal residue on
the first oligopeptide is non glycine then the $N$-
terminal residue on the second oligopeptide is glycine
and if the $N$-terminal residue on the second oligopeptide
is non glycine then the $C$-terminal residue on the first
oligopeptide is glycine. **R'** is selected from the group
consisting of 3-carboxy 4-nitrophenyl and benzyl.

Figure 2 illustrates the chemical synthesis of
$N^\alpha$(substituted) peptide segments **2a** and **2b**.
Figure 3 illustrates the chemical synthesis of compound
**8b**.

- 6 -

Figure 4 shows a table with summaries of model ligations for Examples #1-5. Superscript is as follows: (a) Thioester peptides where -SNB=3-carboxy-4-nitrophenyl thioester and -Sbzl = benzyl thioester; (b) $N^\alpha$(substituted) peptides where etsh=$N^\alpha$(ethanethiol)- and oetsh=$N^\alpha$(oxyethanethiol)-; (c) Based on Peptide **2**, estimated from analytical reverse phase HPLC (peak areas) and ESMS; (d) Adjusted to pH 4.5 after the indicated hrs at pH 7.5.

Figures 5A-5C illustrate the model ligation, Example #2 (infra), to form amide -Gly($N^\alpha$-OCH$_2$CH$_2$SH)Gly- compound **4b** carried out in 6 M guanidine•HCl, 0.1 M Na$_2$HPO$_4$, pH 7.5. (A)-(C) represent HPLC plots as follows: (A) Analytical HPLC (10-50%B over 30 min) at t = 0; peak **a** thioester peptide, **1**, LYRAG-$^\alpha$COSC$_6$H$_3$(3-CO$_2$H-4-NO$_2$); peak **b**, $N^\alpha$(oxyethanethiol) peptide, **2b**, [HSCH$_2$CH$_2$O]-GRNTATIMMQRGNFR-$^\alpha$CONH$_2$; (B) Analytical HPLC (10-50%B over 30 min) at t = 1 h; peak **c**, non-peptide impurity; peak **d**, LYRAG-$^\alpha$COSC$_6$H$_5$, resulting from transthioesterification of **1** (peak **a**) with thiophenol; peak **e**, intermediate ligation product, **4b**, LYRAGG($N^\alpha$-OCH$_2$CH$_2$SH)RNTATIMMQRGNFR-$^\alpha$CONH$_2$ and a minor amount of unreacted [HSCH$_2$CH$_2$O]-GRNTATIMMQRGNFR-$^\alpha$CONH$_2$ (**2b**), as determined by electrospray ionization MS. (C) Analytical HPLC (15-40%B over 30 min) of peak **e** after HPLC-purification and treatment with Zn in acidic HPLC solvent overnight; peak **f**, final ligation product, **5b**, LYRAGGRNTATIMMQRGNFR-$^\alpha$CONH$_2$; peak **g**, non-peptide impurity.

Figures 6A-6C illustrate the model ligation, Example #3 (infra) to form a -Phe($N^\alpha$-OCH$_2$CH$_2$SH)Gly- linked product **4b**. Ligation and rearrangement carried out at 37 °C. (A)-(C) represent HPLC plots as follows: (A) Analytical HPLC (20-40%B over 30 min) after 11.5 h in 6 M guanidine•HCl, 0.1 M Na$_2$HPO$_4$, pH 7.5; peak **a**,

**SUBSTITUTE SHEET (RULE 26)**

$N^\alpha$(oxyethanethiol) peptide, **2b**, [HSCH$_2$CH$_2$O]-
$\underline{G}$RNTATIMMQRGNFR-$^\alpha$CONH$_2$; peak **b**, unrearranged intermediate
ligation product, **3b**, LYRA$\underline{F}$-[$^\alpha$COSCH$_2$CH$_2$O]-
$\underline{G}$RNTATIMMQRGNFR-$^\alpha$CONH$_2$; peak **c**, rearranged intermediate
ligation product, **4b**, LYRA$\underline{FG}$(N$^\alpha$-OCH$_2$CH$_2$SH)RNTATIMMQRGNFR-
$^\alpha$CONH$_2$; peak **d**, thioester peptide, **1**, LYRA$\underline{F}$-$^\alpha$COSCH$_2$C$_6$H$_5$.
(B) Analytical HPLC (20-40%B over 30 min) after 10 h in
6 M guanidine•HCl, 0.1 M NaCH$_3$CO$_2$, pH 4.5.   (C)
Analytical HPLC (0-67%B over 30 min) of HPL$C$-purified
peak **c** after treatment with Zn in acidic HPLC solvent
overnight; peak **f**, final ligation product, **5b**,
LYRA$\underline{FG}$RNTATIMMQRGNFR-$^\alpha$CONH$_2$.


Detailed Description of the Invention

     The invention is directed to a two step method for
chemically ligating together two seperate peptide
fragments.  More particularly, the invention relates to
a general methodology for native (amide-forming)
chemical ligation of unprotected peptide segments to
produce a native peptide bond.


     Step 1 comprises the condensation of a $C$-terminal
thioester of a first peptide with an unoxidized
sulfhydryl moiety of a second oligopeptide and produces
an intermediate oligopeptide which links the first and
second oligopeptides with an δ or γ – amino-thioester
bond.  If the $C$-terminal residue on the first
oligopeptide is non glycine then the N-terminal residue
on the second oligopeptide is glycine and if the $N$-
terminal residue on the second oligopeptide is non
glycine then the $C$-terminal residue on the first
oligopeptide is glycine with the proviso that the non
glycine residue is a non β-branched amino acid.


     Step 2 comprises the rearrangement of the δ or γ –
amino-thioester bond of the intermediate oligopeptide
via intramolecular attack of the δ or γ – amino group

onto the thioester moiety which displaces a sulfhydryl moiety as a byproduct from the thioester moiety and produces an oligopeptide product linking the first and second oligopeptides with a native amide bond. The

5  nitrogen of the amide bond, however, contains the auxiliary functional group with the displaced sulfhydryl moiety.

An optional third step comprises the removal of the

10  auxiliary functional group on the amide nitrogen, from the oligopeptide product with a reducing agent and produces a native peptide bond. The amide nitrogen auxiliary functional group is $N$-α-O-(CH$_2$)$N$-SH wherein $1 \leq n \leq 2$ and the reducing agent is Zinc (underlined

15  nitrogen represents the coupled amide nitrogen).

The method allows for the synthesis of peptides and proteins with either native or modified backbone structures. Figure 1 outlines the procedure. The

20  peptide α-carboxythioester (1) reacts, via thiol exchange, with either an N$^α$(ethanethiol) peptide (2a) or an N$^α$(oxyethanethiol) peptide (2b) to produce the ligation product 3. This thioester-linked intermediate rearranges through a favorable geometric arrangement

25  involving a 5- or 6-membered ring to give the amide-linked product 4, containing a secondary amide in 4a or the analogous $N$-oxyalkyl compound in 4b. The resulting $N$-$^α$(substituted)amides also have potentially favorable solubility properties (Quibell et al. $J. Am. Chem. Soc.$

30  **1995**, $117$, 11656-11668) the $N$-$^α$(O-alkyl)amide ligation product has the further advantage of being stable to HF cleavage conditions, yet is simply removed under mild conditions. In the case of 4b, zinc dust can be added directly to the reverse phase HPLC-purified peptide in

35  the acidic eluant to reduce the $N$-O bond of the O-alkoxyhydroxamate and give the native backbone structure of ligation product 5b.

All peptide segments were synthesized in stepwise
fashion by established solid phase methods using *in situ*
neutralization/2-(1*H*-benzotriazol-1-yl)-1,1,3,3-
tetramethyluronium hexafluorophosphate (HBTU) activation
5     protocols for *tert*-butoxycarbonyl (Boc) chemistry,
purified by preparative reverse phase HPLC, and
characterized by electrospray mass spectrometry (ESMS)
(Schnölzer et al. *Int. J. Pept. Protein Res.* **1992**, *40*,
180-193).   Peptide-α-thioesters (**1**) were generated from
10    the corresponding peptide-α-thioacids, which in turn,
were synthesized on a thioester resin as described in
Canne et al. *Tetrahedron Lett.* **1995**, *36*, 1217-1220).
The peptide-α-thioacids were converted to either the
corresponding 3-carboxy-4-nitrophenyl thioesters in 6 M
15    guanidine•HCl, 0.1 M Na acetate, pH 5.0-6.5, by reaction
with 1.5 equivalents of 5,5'-dithio-bis(2-nitrobenzoic
acid) or to the corresponding benzyl ester in 6M
guanidine•HCl, 0.1 M Na acetate, pH 4.0, using 10
equivalents of benzyl bromide (Dawson et al. *Science*
20    **1994**, *266*, 776-779).   Both these thioesters provide
satisfactory leaving groups for the ligation reactions,
with the 3-carboxy-4-nitrophenyl  thioesters
demonstrating a somewhat faster reaction rate than the
corresponding benzyl thioesters (Dawson et al. *Science*
25    **1994**, *266*, 776-779).

The syntheses of N$^{\alpha}$(ethanethiol)- (**2a**) and
N$^{\alpha}$(oxyethanethiol)- peptides (**2b**) are illustrated in
Figure 2.   The appropriate α-bromocarboxylic acid,
30    activated as the symmetric anhydride (0.5 equivalents
1,3-diisopropylcarbodiimde (DIC) in dichloromethane)
(Robey et al. *Anal. Biochem.* **1989**, *177*, 373-377) was
coupled to the deprotected *N*-terminal amino acid of
peptide-resin **6** to give bromoacyl-peptide-resin **7**.   The
35    bromide was then displaced, with inversion of
stereochemistry, by the amine function of either
structure **8a** or **8b** in DMSO to give peptide-resin **9**.

- 10 -

Deprotection and cleavage from the resin in anhydrous HF
gave **2b** directly or **2a** still in the form of the
disulfide which was reduced to the free thiol.

5      The aminoethanethiol derivative **8a** was synthesized
in one step from the reaction of 2-aminoethanethiol and
5,5'-dithio-bis(2-nitrobenzoic acid) in 80% acetonitrile
in water.  The synthesis of the aminooxyethanethiol
derivative **8b** was more involved and is illustrated in
10     Figure 3.  Bromide **11** was produced from the reaction of
*N*-hydroxyphthalimide (**10**) with a large excess of 1,2-
dibromoethane (Bauer *J. Org. Chem.* **1963**, *28*, 1604-1608).
Bromide **11** was then converted to the protected
aminooxyethanethiol derivative **12** with 4-methylbenzyl
15     mercaptan in the presence of the base 1,8-
Diazabicylco[5.4.0]unde*C*-7-ene (DBU).  The phthalimide
group of **12** was removed in a two step process (Osby  et
al. *Tetrahedron Lett.* **1984**, *25*, 2093-2096) involving
reduction with $NaBH_4$ followed by treatment with acetic
20     acid to give the desired aminooxyethanethiol derivative
**8b**.

       Compared with the original native chemical ligation
at **X**-Cys sites (Dawson et al. *Science* **1994**, *266*, 776-
25     779), the most notable feature of the ligation chemistry
of this invention is the slower rearrangement of the
initial thioester ligation product.  Intramolecular
attack by the $\alpha$-NH(O-alkyl) group via a six-membered
ring intermediate in the current instance is
30     considerably less favored (Mandolini et al. *J. Am. Chem.
Soc.* **1978**, *100*, 550-554) than the facile five-membered
ring mediated attack of the unsubstituted $\alpha$-NH$_2$ group in
the **X**-Cys ligation (Dawson et al. *Science* **1994**, *266*,
776-779).  Except for the Gly-Gly ligation, one is able
35     to isolate the thioester-linked initial ligation product
of the present invention (Figure 4), which rearranged
only slowly.  Remarkably, for $C^\alpha$-substituted amino acids

●                                    ●

on both sides of the ligation site, one is unable to observe susbsequent rearrangement to an amide bond:  the thioester-linked intermediate was indefinitely stable under the conditions of reaction.  These slow
5    rearrangements are similar to those observed in an (*N*-substituted)amide-forming ligation chemistry previously described (Liu et al. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 6584-6588).

10        The use, in the invention, of a temporary auxillary functional group in amide-forming ligation is reminiscent of the 'thiol capture' strategy proposed by Kemp (Fotouhi et al. *J. Org. Chem.* **1989**, *54*, 2806-2817).  Both native chemical ligation and the thiol-
15   capture method have as their stated goal the use of unprotected peptides in a segment condensation strategy to achieve the synthesis of long polypeptide chains. The key aspect of the native chemical ligation approach is reaction under conditions that promote exchange of
20   the thiol moiety of the initial thioester-linked intermediate products, to give regioselective ligation at the *N*-terminal Cys even in the presence of other Cys residues in both segments (Dawson et al. *Science* **1994**, *266*, 776-779).  By contrast, in the 'thiol capture'
25   strategy and related methods (Liu et al. *Tetrahedron Lett.* **1996**, *37*, 933-936) regiospecific protection of Cys side chain functionalities prior to the ligation reaction is a necessity (Kemp et al. *Peptides: Proceedings of the 11th American Peptide Symposium*
30   Rivier, J.E.; Marshall, G.R., Eds, ESCOM:  Lieden, 1990; pp 920-922), thus frustrating their original intent.

         Similarly, the use, in the current invention, of a carbonyl-activated α-thiocarboxyl function (i.e. an α-
35   COSR) is reminiscent of chemistry previously used in peptide segment condensation (Blake et al. *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 4055-4058; Yamashiro et al.

*Int. J. Pept. Protein Res.* **1988**, *31*, 322-334; Hojo et al. *Peptide Chemistry* Okada, Y., Ed, Protein Research Foundation: Osaka, 1994; pp 9-12) However, these syntheses were based on conventional no*N*-chemoselective attack by the α-amine nucleophile of the second segment on a Ag⁺-activated peptide-ᵅCOSR, and thus necessitated regiospecific (re)protection of all other α- and ɛ-amine functional groups in both segments.

By contrast, thioester-mediated amide-forming ligation chemistry is compatible with the use of completely unprotected peptide segments with the full range of side chain functionalities found in nature, including thiols (Baca et al. *J. Am. Chem. Soc.* **1995**, *117*, 1881-1887; Dawson et al. *Science* **1994**, *266*, 776-779). For this reason, native chemical ligation is simple, practical, and is a general approach to the total chemical synthesis of proteins provided they contain appropriate ligation sites.

This invention increases the utility of the native chemical ligation method reported by Dawson, *et al* (Dawson et al. *Science* **1994**, *266*, 776-779) by extending the number of dipeptide sequences that can be used as ligation sites. In addition to the **X**-Cys ligation site of that methodology, it is disclosed herein how make use of **X**-Gly and Gly-**X** ligation sites (i.e. the *C*-terminal residue being glycine if the *N*-terminal residue is non-glycine or conversely, the *N*-terminal residue being glycine if the *C*-terminal residue is non-glycine). In the preferred mode, **X** is any of the non-ß-branched amino acids. However, our previous studies have shown that **X** can be any amino acid, including ß-branched such as Val;(Dawson et al. *Science* **1994**, *266*, 776-779).

The invention potentially extends the number of suitable sites for native chemical ligation by a factor

of three, to more than 50 of the 400 dipeptide sequences found in proteins. Because there is considerable latitude in choosing a ligation site in a target sequence which contains a glycine residue, this extended applicability will render most polypeptides accessible by native chemical ligation.

Native chemical ligation in its original form (Dawson et al. *Science* **1994**, *266*, 776-779) or in the form described here provides direct synthetic access to polypeptide chains the size of typical protein domains. Other ligation chemistries (Schnölzer et al. *Science* **1992**, *256*, 221-225; Rose et al. *Am. Chem. Soc.* **1994**, *116*, 30-34; Liu et al. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 6584-6588) can be used to join synthetic domains in a modular fashion to produce large (i.e. >20 kilodalton), fully functional synthetic proteins (Canne et al. *J. Am. Chem. Soc.* **1995**, *117*, 2998-3007; Baca et al. *J. Am. Chem. Soc.* **1995**, *117*, 1881-1887). In its most general form, incorporating all suitable chemistries. (Dawson et al. *Science* **1994**, *266*, 776-779) the chemical ligation approach represents the next stage in the evolution of methods for the chemical synthesis of polypeptides and, for the first time, provides for repoducible, practical total chemical synthesis of proteins.

## Example Ligations 1-4

Results of example ligations are summarized in Figure 4. All ligations were run at concentrations ranging from 4 mgs/mL to 8 mgs/mL of each peptide in either 8M Urea, 0.1 M $Na_2HPO_4$, pH 7.0 or 6M guanidine•HCl, 0.1 M $Na_2HPO_4$, pH 7.5. Immediately after solvation of the peptide segments, 2-5% benzyl mercaptan (example ligation #1) or thiophenol (example ligations #2-5), by volume of ligation buffer, was added to keep thiol functions in the reduced form. This had the added

● ●

consequence of exchanging a significant amount of the original peptide-$\alpha$thioester (**1**) to a peptide-$\alpha$thiobenzyl ester or peptide-$\alpha$thiophenyl ester, both of which are still capable of reacting in the desired fashion with

5　peptide **2**. The ligation reactions were followed by analytical reverse-phase HPLC and the products identified by ESMS. In the following examples, the amino acid residues involved in the ligation are underlined.

10

## Example 1

The peptide-$\alpha$thio(3-carboxy-4-nitro)phenyl ester **1** had the sequence LYRA<u>G</u>-$\alpha$COSC$_6$H$_3$(3-CO$_2$H-4-NO$_2$) (SEQ ID NO 1) (observed mass 760 Da, calcd 760 Da). The

15　N$^\alpha$(ethanethiol) peptide **2a**, protected as the 5-thio-2-(nitrobenzoic acid) disulfide, had the sequence [(3-CO$_2$H-4-NO$_2$)-C$_6$H$_3$-S-SCH$_2$CH$_2$]-<u>G</u>AGPAGD-$\alpha$CONH$_2$ (SEQ ID NO 2) (observed mass 800 Da, calcd 800 Da) which was reduced to [HSCH$_2$CH$_2$]-<u>G</u>AGPAGD-$\alpha$CONH$_2$ (SEQ ID NO 2) (observed mass

20　603 Da, calcd 603 Da) in the ligation mixture. After one hour, the product of the ligation, **4a**, [LYRA<u>GG</u>(N$^\alpha$-CH$_2$CH$_2$SH)AGPAGD-$\alpha$CONH$_2$ (SEQ ID NO 3) (observed mass 1163±1 Da, calcd 1163 Da)] had already formed to a significant degree. After purification of the ligation product by

25　reverse phase HPLC, it was treated with 5,5'-dithio-bis(2-nitrobenzoic acid) for one hour in 6 M urea, 0.1 M phosphate, pH 6.0 to yield the expected -S-nitrobenzoic acid disulfide product, LYRA<u>GG</u>[N$^\alpha$-CH$_2$CH$_2$S-SC$_6$H$_3$(3-CO$_2$H-4-NO$_2$)]AGPAGD-$\alpha$CONH$_2$ (SEQ ID NO 3) (observed mass 1360±1

30　Da, calcd 1360 Da). This confirmed the rearrangement of **3a** to form **4a**, since **3a** would have been unreactive towards 5,5'-dithio-bis(2-nitrobenzoic acid). The results are summarized in Figure 4.

35　## Example 2

Figures 5A-5C show the course of example ligation #2. The peptide-$\alpha$thioester segment **1** (peak **a**) had the

sequence LYRAG-$^{\alpha}$COSC$_6$H$_3$(3-CO$_2$H,4-NO$_2$) (SEQ ID NO 1)
(observed mass 760 Da, calcd 760 Da). The
N$^{\alpha}$(oxyethanethiol) peptide **2b** (peak **b**) consisted of the
sequence [HSCH$_2$CH$_2$O]-GRNTATIMMQRGNFR-$^{\alpha}$CONH$_2$ (SEQ ID NO 4)

5    (observed mass 1827±1 Da, calcd 1828 Da). After one
hour reaction at room temperature in the presence of
added thiophenol, the ligation product, LYRAGG(N$^{\alpha}$-
OCH$_2$CH$_2$SH)RNTATIMMQRGNFR-$^{\alpha}$CONH$_2$ (SEQ ID NO 5) (**4b**,
observed mass 2388±1 Da, calcd 2389 Da) had formed to a

10   significant degree, coeluting with unreacted
N$^{\alpha}$(oxyethanethiol) peptide **2b** (peak **e**). After further
reaction overnight at room temperature, the ligation
product was purified by HPLC and zinc dust was added
directly to the collected peptide in HPLC eluant and

15   stirred overnight at room temperature. Under these
conditions, the zinc effectively reduced the *N*-O bond.
Reductions of this type are possible through a variety
of reagents (Keck et al. *Tetrahedron Lett.* **1995**, *36*,
7419-7922). The resulting peptide (peak **f**) gave a mass

20   consistent with reduction of the *N*-O bond to produce
LYRAGGRNTATIMMQRGNFR-$^{\alpha}$CONH$_2$ (SEQ ID NO 5) (**5b**, observed
mass 2313±1 Da, calcd 2313 Da). This confirmed the
rearrangement of **3b** to **4b** as shown in Figure 1, since
cleavage of the *N*-O bond in unrearranged **3b** would have

25   resulted in the formation of two distinct peptides of
significantly lower masses.


Example 3

       Figures 6A-6C show the course of example ligation
30   #3. The peptide-$^{\alpha}$thioester segment **1** (peak **d**) consisted
of a peptide with a *C*-terminal phenylalanine thioester
[LYRAF-$^{\alpha}$COSCH$_2$C$_6$H$_5$ (SEQ ID NO 6) (observed mass 775 Da,
calcd 775 Da)], thus providing a more sterically
hindered model than the previous cases (example

35   ligations #1 and #2) which involved *C*-terminal glycine
thioesters. The N$^{\alpha}$(oxyethanethiol) peptide **2b** (peak **a**)
consisted of the sequence [HSCH$_2$CH$_2$O]-GRNTATIMMQRGNFR-

- 16 -

$^{\alpha}$CONH$_2$ (SEQ ID NO 7) (observed mass 1827±1 Da, calcd 1828
Da). The presence of the more sterically hindered
thioester slowed the reaction relative to the unhindered
models. However, heating at 37 °C was found to
accelerate the rate of initial ligation. Figure 6A
illustrates the ligation reaction after 11.5 hours at 37
°C. The rate of rearrangement of **3b** to **4b** was slowed
enough to observe the unrearranged product **3b**, LYRA_F_-
[$^{\alpha}$COSCH$_2$CH$_2$O]-_GRNTATIMMQRGNFR_-$^{\alpha}$CONH$_2$ (SEQ ID NO 8) (peak
**b**, observed mass 2478±1 Da, calcd 2478 Da), eluting
slightly before the rearranged product **4b**, LYRA_FG_(N$^{\alpha}$-
OCH$_2$CH$_2$SH)RNTATIMMQRGNFR-$^{\alpha}$CONH$_2$ (SEQ ID NO 9) (peak **c**,
observed mass 2478±1 Da, calcd 2478 Da). Having
identical masses, the unrearranged and rearranged
products, **3b** and **4b**, were identified by their zinc
reduction products. Intermediate **3b**, peak **b**, gave two
peptides upon zinc reduction with masses of 669 Da and
1753 Da corresponding to the peptide sequences LYRA_F_-
$^{\alpha}$COOH (thioester hydrolysis) and _GRNTATIMMQRGNFR_-$^{\alpha}$CONH$_2$,
respectively. It was subsequently determined that
lowering the pH to 4.5 after initial ligation by
diluting the crude reaction mixture (to five times the
volume) with 6M guanidine•HCl, 0.1 M Na acetate, pH 4.0
accelerated the rate of rearrangement. Figure 6B shows
that rearrangement of the initial ligation product **3b** to
**4b** (peak **c**) was complete after 10 hours at pH 4.5 at
37°C. HPLC purification of this peak and subsequent
zinc reduction gave a peptide (peak **f**) of the expected
mass, LYRA_FG_RNTATIMMQRGNFR-$^{\alpha}$CONH$_2$ (SEQ ID NO 9) (**5b**,
observed mass 2403±1 Da, calcd 2404 Da).

Example 4

LYRA_G_-$^{\alpha}$COSCH$_2$C$_6$H$_5$ **1** (SEQ ID NO 1) (observed mass 685
Da, calcd 685 Da) was ligated to [HSCH$_2$CH$_2$O]-
_A_ARHTVHQRHLHG-$^{\alpha}$COOH **2b** (SEQ ID NO 10) (observed mass
1595±1 Da, calcd 1596 Da) at pH 7.5. This example
provided steric hindrance, in the form of an Ala

residue, on the opposite side of the ligation site from example ligation **#3**. The rate of reaction was similiar to that of example 3, the rate being significantly slower than the unhindered examples (example ligations **#1** and **#2**), but enhanced with heating at 37 °C. Both unrearranged (**3b**, LYRAG-$^\alpha$[COSCH$_2$CH$_2$O]-AARHTVHQRHLHG-$^\alpha$COOH) (SEQ ID NO 11) and rearranged (**4b**, LYRAGA(N$^\alpha$-OCH$_2$CH$_2$SH)ARHTVHQRHLHG-$^\alpha$COOH) (SEQ ID NO 12) ligation products were formed, observed mass (both unrearranged and rearranged) 2156±1 Da, calcd 2157 Da. However, unlike example ligation **#3**, the rate of rearrangement was **not** enhanced by lowering the pH to 4.5 after initial ligation, though rearrangment was nearly complete after 2 days at 37 °C regardless of pH as shown by successful reduction to the amide. It should also be noted that the presence of a significant number of histidines in the final ligation product **5b** (LYRAGAARHTVHQRHLHG-$^\alpha$COOH (SEQ ID NO 12), observed mass 2080±1 Da, calcd 2080 Da) resulted in binding of the peptide to the zinc. EDTA had to be added to the HPLC buffer/Zn mixture to free the peptide from the zinc after reduction of the *N-O* bond. The results are summarized in Figure 4.

Example 5

Example 5 provides an example with steric bulk on both sides of the ligation site. LYRAF-$^\alpha$COSCH$_2$C$_6$H$_5$ **1** (SEQ ID NO 6) (observed mass 775 Da, calcd 775 Da) was ligated to [HSCH$_2$CH$_2$O]-AARHTVHQRHLHG-$^\alpha$COOH **2b** (SEQ ID NO 13) (observed mass 1595±1 Da, calcd 1596 Da) at pH 7.5. Though the initial (i.e. unrearranged) ligation product **3b** was observed, LYRAF-$^\alpha$[COSCH$_2$CH$_2$O]-AARHTVHQRHLHG-$^\alpha$COOH (SEQ ID NO 14) (observed mass 2246±1 Da, calcd 2247 Da), there was no evidence of rearrangement over time, even at lower pH. The presence of side chains on both sides of the ligation site apparently provided too much steric hindrance for the rearrangement to occur via a 6-membered ring intermediate under the conditions used.

- 18 -

The results are summarized in Figure 4.

The above examples emphasize the extraordinary
facility of our original native ligation chemistry
5    (Dawson et al. *Science* **1994**, *266*, 776-779) which
recently was independently repeated in essentially
identical form in model ligation studies (Tam et al.
*Proc. Natl. Acad. Sci. USA* **1995**, *92*, 12485-12489). We
have used this original native peptide bond-forming
10    ligation reaction in the chemical synthesis of a number
of proteins with full biological activity, including the
chemokine IL-8 (Dawson et al. *Science* **1994**, *266*, 776-
779), the enzymes HIV-1 protease and barnase, the serine
proteinase inhibitors turkey ovomucoid third domain and
15    eglin C, and a b/HLH transcription factor. This ligation
reaction was based on principles enunciated by Max
Brenner (Brenner M. *Peptides. Proceedings of the Eighth
European Peptide Symposium* Beyerman, H.C., Ed., North
Holland: Amsterdam, 1976; pp 1-7) and made use of
20    chemistry first described by Wieland (Wieland et al.
*Liebigs Ann. Chem.* **1953**, *583*, 129-149).

### EXPERIMENTAL PROTOCALS

25    <u>General</u>

Machine-assisted solid-phase peptide syntheses were
carried out on a custom-modified Applied Biosystems 430A
peptide synthesizer (Schnölzer et al. *Int. J. Pept.
Protein Res.* **1992**, *40*, 180-193). Reverse-phase HPLC was
30    performed on a Rainin HPLC system with 214-nm UV
detection, using Vydac *C*-18 analytical (5 $\mu$m, 0.46 X 15
cm) and semipreparative (10 $\mu$m, 1.0 X 25cm) columns.
Chromatographic separations were achieved using linear
gradients of buffer B in A (A = 0.1% TFA in water, B=
35    90% $CH_3CN$/10%water containing 0.09% TFA) over 30-60 min
at 1 mL/min (analytical) or 3 mL/min (semipreparative).
Mass spectra of all peptide segments were obtained with

a Sciex API-III electrospray quadrupole mass
spectrometer; observed masses were derived from the
experimental m/z values for all observed protonation
states of a molecular species, using the program MacSpec
(Sciex). Calculated masses were based on average
isotope composition and were derived using the program
MacProMass (Terry Lee and Sunil Vemuri, Beckman Research
Institute, Duarte, CA). All other mass spectrometry was
performed at The Scripps Research Institute Mass
Spectrometry Facility. $^1$H NMR spectra were recorded on a
Bruker 250 MHz spectrophotometer and chemical shifts are
reported in parts per million downfield from $Me_4Si$.
Microanalyses were performed at The Scripps Research
Institute X-ray Crystallographic Facility and agreed
with calculated values ±0.4%. BoC-L-amino acids and
HBTU were purchased from Novabiochem (La Jolla, CA). 4-
Hydromethylphenylacetamidomethyl (PAM) resins and
diisopropylethylamine (DIEA) were obtained from Applied
Biosystems (Foster City, CA), and methylbenzhydrylamine
(MBHA) resin was obtained from Peninsula Laboratories,
Inc (San Carlos, CA). Synthesis grade dimethylformamide
(DMF) was obtained from Baker, and AR grade $CH_2Cl_2$ and
HPLC-grade $CH_3CN$ obtained from Fisher. Trifluoroacetic
acid (TFA) was obtained from Halocarbon (New Jersey).
HF was purchased from Matheson Gas. 4-Methylbenzyl
mercaptan was obtained from Lancaster. All other
reagents were AR grade or better and were obtained from
Aldrich Chemical or from Fisher.

## Peptide Segment Synthesis Chain Assembly

Peptides were synthesized in stepwise fashion by
established machine-assisted or manual solid-phase
methods using *in situ* neutralization/HBTU activation
protocols for Boc chemistry (Schnölzer et al. *Int. J.
Pept. Protein Res.* **1992**, *40*, 180-193). Side chain
protection was as follows: BoC-Arg(*p*-toluenesulfonyl)-
OH, BoC-Asn(xanthyl)-OH, BoC-Asp(O-cyclohexyl)-OH, BoC-

His(dinitrophenyl)-OH, Bo$C$-Thr(benzyl)-OH, and Bo$C$-
Tyr(2-bromobenzyloxycarbonyl)-OH.  Bo$C$-Gl$N$-OH and Bo$C$-
Met-OH were used without side chain protection.
Coupling reactions were monitored by quantitative
5    ninhydrin assay (Sarin et al *Anal. Biochem.* **1981**, *117*,
147-157) and were typically >99%.  After chain assembly
was complete, peptides were deprotected and
simultaneously cleaved from the resin by treatment with
HF containing 5% *p*-cresol for 1 h at 0 °C to give the
10    peptide-$^\alpha$COSH, -$^\alpha$CONH$_2$, or -$^\alpha$CO$_2$H.  After removal of the
HF under reduced pressure, the crude peptide was
precipitated in anhydrous Et$_2$O (diethyl ether), dissolved
in HPLC buffer (40-50% B) and lyophilized.


15    <u>Synthesis of Peptide-$^\alpha$thioesters (1) as shown in Figure 1</u>
     Thioacid peptides were synthesized on the
appropriate Bo$C$-aminoacyl-S-Resins, made by coupling [4-
[$\alpha$-($N$-$t$-Bo$C$-aminoacyl-S)benzyl]phenoxy]acetic acid, DCHA
salt (Canne et al. *Tetrahedron Lett.* **1995**, *36*, 1217-
20    1220), (2.0 equiv) and aminomethyl-resin (1 equiv,
washed with 10% DIEA in DMF; Aldrich) with HBTU (1.6
equiv; Aldrich) added as an activating agent and DIEA (1
equiv; Aldrich), in DMF (dimethylformamide).  Peptide-
$^\alpha$COSC$_6$H$_3$(3-CO$_2$H-4-NO$_2$) thioesters were generated by
25    dissolving the crude peptide-$^\alpha$COSH (15-20 mgs) in 6 M
guanidine•HCl, 0.1 M Na Acetate, pH 5.0-6.5, to which
was added 1.5 equivalents of 5,5'-dithio-bis(2-
nitrobenzoic acid) (Dawson et al. *Science* **1994**, *266*,
776-779).  The mixture was vortexed briefly and
30    purified after 10 min.  Identity of the peptide-
$^\alpha$COSC$_6$H$_3$(3-CO$_2$H-4-NO$_2$) ester was unambigously confirmed by
precise electrospray mass measurements, in contrast with
Liu et al. *Tetrahedron Lett.* **1996**, *37*, 933-936.
Peptide-$^\alpha$COSCH$_2$C$_6$H$_5$ thioesters were generated by
35    dissolving the crude peptide-$^\alpha$COSH (15-20 mgs) in 6M
guanidine•HCl, 0.1 M Na acetate, pH 4.0, to which was
added 10 equivalents of benzyl bromide (Dawson et al.

*Science* **1994**, *266*, 776-779). The mixture was vortexed briefly and purified after 1 h. LYRAG-$^\alpha$COSC$_6$H$_3$(3-CO$_2$H-4-NO$_2$) (observed mass 760 Da, calcd 760 Da) was purified by semipreparative HPLC (20-60%B over 40 min) to give 20-30% yield. LYRAG-$^\alpha$COSCH$_2$C$_6$H$_5$ (observed mass 685 Da, calcd 685 Da) was purified by semipreparative HPLC (15-45%B over 60 min) to give 25-30% yield. LYRAF-$^\alpha$COSCH$_2$C$_6$H$_5$ (observed mass 775 Da, calcd 775 Da) was purified by semipreparative HPLC (30-60%B over 60 min) to give 25-30% yield. Most of the loses in yield arose simply from HPLC recoveries.

## Synthesis of N$^\alpha$(ethanethiol) and N$^\alpha$(oxyethanethiol) peptides (2) as illustrated in Figure 2

These peptides were synthesized on either MBHA (Sigma)or the appropriate Bo$C$-aminoacyl-OCH$_2$-PAM-resins (Sigma). After chain assembly was complete and the N$^\alpha$Boc group removed with neat TFA (two 1 min treatments) and neutralized with 10% DIEA in DMF (two 1 min treatments), the peptide was bromoacetylated by the method of Robey (Robey et al. *Anal. Biochem.* **1989**, *177*, 373-377). Bromoacetic acid (2.0 mmol) was dissolved in CH$_2$Cl$_2$ (2 mL) to which was added DIC (1 mmol; 2-dimethylaminoisopropyl chloride hydrochloride; Aldrich or Sigma). After activation for 10-15 min to form the symmetric anhydride, the mixture was diluted with DMF (2 mL), added to the peptide-resin and coupled for 30 min. The resin was then washed with DMSO and **8a** (2 M in DMSO) or **8b** (1 M in DMSO) was added and allowed to react with the bromoacetylated peptide-resin for 8-23 h. The peptides were purified without further modifcation after cleavage from the resin. [(3-CO$_2$H-4-NO$_2$)-C$_6$H$_3$-S-SCH$_2$CH$_2$]-AGPAGD-$^\alpha$CONH$_2$ (**2a**)(observed mass 800 Da, calcd 800 Da) was purified by semipreparative HPLC (0-67%B over 60 min) to give ⁻24% yield. [HSCH$_2$CH$_2$O]-GRNTATIMMQRGNFR-$^\alpha$CONH$_2$ (**2b**)(observed mass 1827±1 Da, calcd 1828 Da) was purified by semipreparative HPLC (15-40%B over 60 min)

to give ~25% yield.    [HSCH$_2$CH$_2$O]-AARHTVHQRHLHG-αCOOH (2b) was synthesized by the above method using the racemate of 2-bromoproprionic acid instead of bromoacetic acid.   The resulting crude lyophilized product was a mixture of the desired peptide and the peptide that results from elimination of HBr from the bromopropionylated peptide (CH$_2$=CHCO-ARHTVHQRHLHG-$^{\alpha}$COOH, observed mass 1502±1 Da, calcd 1502 Da). The mixture was purified by semipreparative HPLC (10-40%B over 60 min) to give [HSCH$_2$CH$_2$O]-AARHTVHQRHLHG-αCOOH (observed mass 1595±1 Da, calcd 1596 Da) in ~10% yield.

## Synthesis of S-[(3-Carboxy-4-nitro)phenylthio]-2-aminoethanethiol (8a) as shown in Figure 2

2-Aminoethanethiol (1.0 g, 13.0 mmol) and 5,5'-dithio-bis(2-nitrobenzoic acid) (1.75 g, 4.4 mmol) were combined in acetonitrile (100 mL) and water (25 mL) and stirred at room temperature for 14 hr.  The reaction mixture was diluted with water (450 mL) and purified by preparative reverse phase HPLC on a Waters Delta Prep 4000 with a Vydac 5.0 x 2.5 cm preparative $C$-18 column to give 8a (1.0 g, 40%)  $^{1}$H NMR (D$_2$O):  δ 7.95 (d, 1H, J=8.7 Hz), 7.61 (dd, 1H, J=8.7, 2.2 Hz),  7.58 (d, 1H, J=2.2 Hz), 3.18 (t, 2H, J=6.5 Hz), 2.92 (t, 2H, J=6.7 Hz).

## Synthesis of N-(2-Bromoethoxy)phthalimide (11) as illustrated in Figure 3.

N-(2-Bromoethoxy)phthalimide (11) was synthesized by a modification of the procedure of Bauer and Sureshl (Bauer et al. J. Org. Chem. 1963, 28, 1604-1608).  N-Hydroxyphthalimide (16.0 g, 98 mmol; Aldrich), triethylamine (30 mL, 215 mmol), and 1,2-dibromoethane (35 mL, 406 mmol; Aldrich) were combined in DMF (115 mL) and stirred at room temperature overnight.  Solids were filtered and washed with DMF.  The filtrate was diluted with water (800 mL) and the resulting precipitate

- 23 -

filtered, dissolved in EtOAc (200 mL) and washed with 1
N HCl (2 x 50 mL), water (1 x 50 mL), saturated NaCl (1
x 50 mL) and dried over $MgSO_4$. Volatiles were removed *in
vacuo* and the resulting solid was recrystallized from

5   95% EtOH to give **11** as a white solid (16.6 g, 63%). $^1$H
NMR ($CDCl_3$): δ 7.82 (m, 4H), 4.49 (t, 2H, J=6.9 Hz),
3.65 (t, 2H, J=6.9 Hz); FAB MS (sodium ion): calc for
$[C_{10}H_8BrNO_3, H^+]$ 291.9585, found 291.9579.


10   <u>Synthesis of *N*-[2-[S-(4-Methylbenzyl)]mercapto]ethoxy]
phthalimide (12) as illustrated in Figure 3</u>

Bromide **11** (16.6, 62 mmol), 4-methylbenzyl mercaptan
(8.5 mL, 63 mmol), and DBU (9.5 mL, 64 mmol; 1,8-
diazabicyclo[5.4.0]unde*C*-7-ene; Aldrich) were combined

15   in benzene (150 mL) and stirred at room temperature for
8 hr. Solids were filtered, washed with benzene and the
filtrate washed with 1 N HCl (2 x 35 mL), water (1 x 35
mL), saturated NaCl (1 x 35 mL) and dried over $MgSO_4$.
Volatiles were removed *in vacuo* and the resulting solid

20   recrystallized from EtOAc/hexane to yield **12** as a white
solid (14.8 g, 74%) $^1$H NMR ($CDCl_3$): δ 7.80 (m, 4H),
7.18 (d, 2H, J=8.0 Hz), 7.04 (d, 2H, J=8.0 Hz), 4.22
(t, 2H, J=7.4 Hz), 3.75 (s, 2H), 2.79 (t, 2H, J=7.4 Hz),
2.27 (s, 3H); FAB MS: calc for $[C_{18}H_{17}NO_3S, H^+]$ 328.1007,

25   found 328.1016. Anal. Calcd for $C_{18}H_{17}NO_3S$ C, 66.03; H,
5.23; N, 4.28; S, 9.79. Found C, 66.04; H, 4.95; N,
4.30; S, 9.58.


<u>Synthesis of S-(4-Methylbenzyl)-2-aminooxyethanethiol
30   (8b) as illustrated in Figure 3</u>

S-(4-methylbenzyl)-2-aminooxyethanethiol was
synthesized by the method of Osby (Osby et al.
*Tetrahedron Lett*. **1984**, *25*, 2093-2096). The *N*-
substituted phthalimide **12** (7.4 g, 23 mmol) was

35   suspended in isopropanol (203 mL) and water (35 mL) to
which was added $NaBH_4$ (3.5 g, 92 mmol). The mixture was
stirred at room temperature overnight. Acetic acid (25

mL) was slowly added till bubbling ceased, the flask
stoppered and heated to 50 °C for 2-3 hr.  Volatiles
were removed *in vacuo*, the resulting solution diluted
with 1 N NaOH and extracted with EtOAc (4 x 50 mL).  The
5   combined EtOAc extractions were washed with saturated
NaCl (1 x 50 mL) and dried over $MgSO_4$.  Volatiles were
removed *in vacuo* and the resulting oil purified by flash
chromatography (1:1 hexane:EtOAc) to yield **8b** as a
clear, colorless oil (3.2 g, 72%) $^1$H NMR ($CDCl_3$):  δ
10  7.21 (d, 2H, J=8.0 Hz), 7.12 (d, 2H, J=8.0 Hz),  5.40
(br s, 2H, $D_2O$ ex.), 3.77 (t, 2H, J=6.5 Hz), 3.71 (s,
2H), 2.64 (t, 2H, J=6.5 Hz), 2.33 (s, 3H); FAB MS:  calc
for [$C_{10}H_{15}NOS$, $H^+$] 198.0953, found 198.0958.  Anal.
Calcd for $C_{10}H_{15}NOS$  C, 60.88; H, 7.66; N, 7.10; S, 16.25.
15  Found  C, 60.79; H, 7.88; N, 7.03; S, 16.11.

- 25 -

What is claimed is:

1.  A method for ligating a first oligopeptide with a
second oligopeptide end to end for producing an
oligopeptide product, the method comprising the
following steps:

  Step A:  admixing the first and second oligopeptides in
    a reaction solution including a catalytic thiol, the
    first oligopeptide including a $C$-terminal residue
    having a $C$-terminal thioester, the second
    oligopeptide including an $N$-terminal residue having
    an $N$-terminal auxiliary functional group with an
    unoxidized sulfhydryl moiety, the $C$-terminal residue
    being glycine if the $N$-terminal residue is non-
    glycine, the $N$-terminal residue being glycine if the
    $C$-terminal residue is non-glycine;

  Step B:  condensing the unoxidized sulfhydryl moiety of
    the $N$-terminal auxiliary functional group with the
    $C$-terminal thioester for producing an intermediate
    oligopeptide linking the first and second
    oligopeptides with an aminothioester bond; and then

  Step C:  rearranging the aminothioester bond of the
    intermediate oligopeptide of said Step B for
    producing the oligopeptide product linking the first
    and second oligopeptides with an amide bond having a
    $N$-linked auxiliary functional group.

2.  A method as described in Claim 1 wherein, the
catalytic thiol is selected from the group consisting of
benzyl mercaptan and thiophenol.

3.  A method as described in claim 2 wherein the $N$-
terminal auxiliary functional group is $N-\alpha-(CH_2)N-SH$
where $1 \leq n \leq 3$.

4.  A method as described in claim 2 comprising the
following additional step:

Step D: Removing the *N*-terminal auxiliary functional group from the oligopeptide product of said step C with a reducing agent for producing a native peptide bond.

5    5.  A method as described in claim 4 wherein the *N*-terminal auxiliary functional group is is $N-\alpha-O-(CH_2)N-SH$ where $1 \leq n \leq 2$.

6.  A method as described in claim 5 wherein the
10   reducing agent is Zinc.

7.  An oligopeptide intermediate comprising:

a first oligopeptide segment including a *C*-terminal thioester,

15   a second oligopeptide segment including a *N*-terminal auxiliary functional group having an unoxidized sulfhydryl moiety, and

an aminothioester linkage unit linking the *C*-terminal thioester and the sulfhydryl group of the

20   auxiliary functional group, said aminothioester linkage unit spontaneously rearranging intramolecularly to form an amide bond linking said first and second oligopeptides segments end to end.

25   8.  A linkage unit for joining two peptide sequences, i.e. an amino terminal peptide sequence and a carboxyl terminal peptide sequence, the linkage unit comprising:

a dipeptide having a first amino acid residue (aa$_1$), a second amino acid resiude (aa$_2$), and an exploded
30   linkage between said first and second amino acid residues,

said first amino acid residue having a backbone carbonyl group, i.e. (aa$_1$-CO-),

said second amino acid residue having a backbone
35   amino group, i.e. (-NR-aa$_2$), wherein R is an auxiliary moiety selected from the group consisting of $\alpha-O-(CH_2)N$-SH and $\alpha-(CH_2)N$-SH, wherein $1 \leq n \leq 2$,

said carbonyl group on the first amino acid residue
having a single bond linking said second amino acid
residue to the first amino acid at the sulfydryl group
of said auxiliary moiety of the second amino acid,

5      said exploded linkage between said first and said
second amino acid residue being represented by the
formula

$(aa_1-CO)-S-(CH_2)N-X-(N-aa_2)$ wherein $1 \leq n \leq 2$ and **X** is
selected from the group consisting of oxygen and
10     methylene,

whereby said first amino acid residue ($aa_1$) of said
dipeptide being employable for forming a peptide linkage
with said amino terminal peptide sequence and said
second amino acid residue ($aa_2$) of said dipeptide being
15     employable for forming a peptide linkage with said
carboxyl terminal peptide sequence, thereby linking said
amino terminal peptide sequence to said carboxyl
terminal peptide sequence with said exploded linkage.

●                    ●
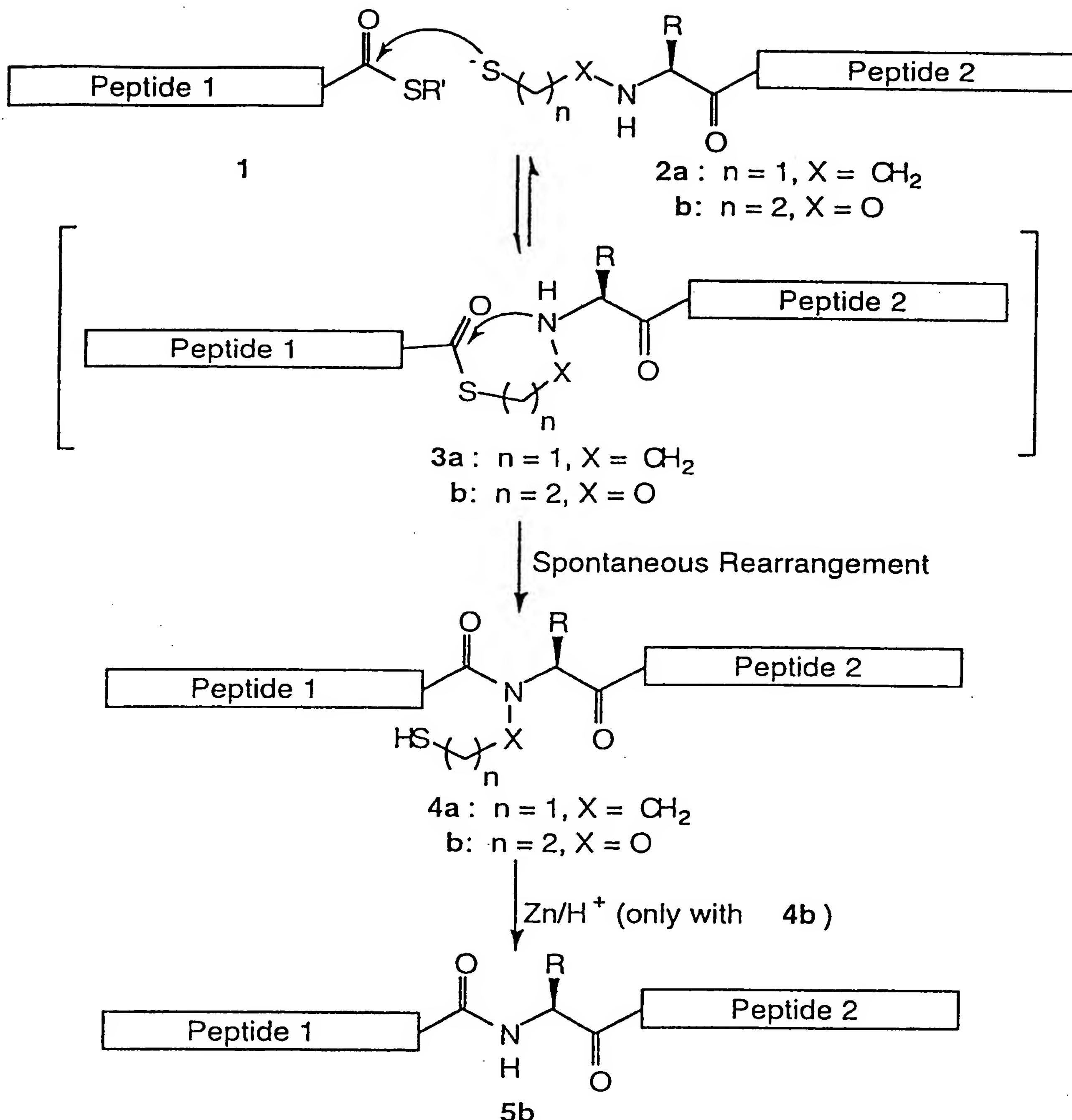
FIG. I
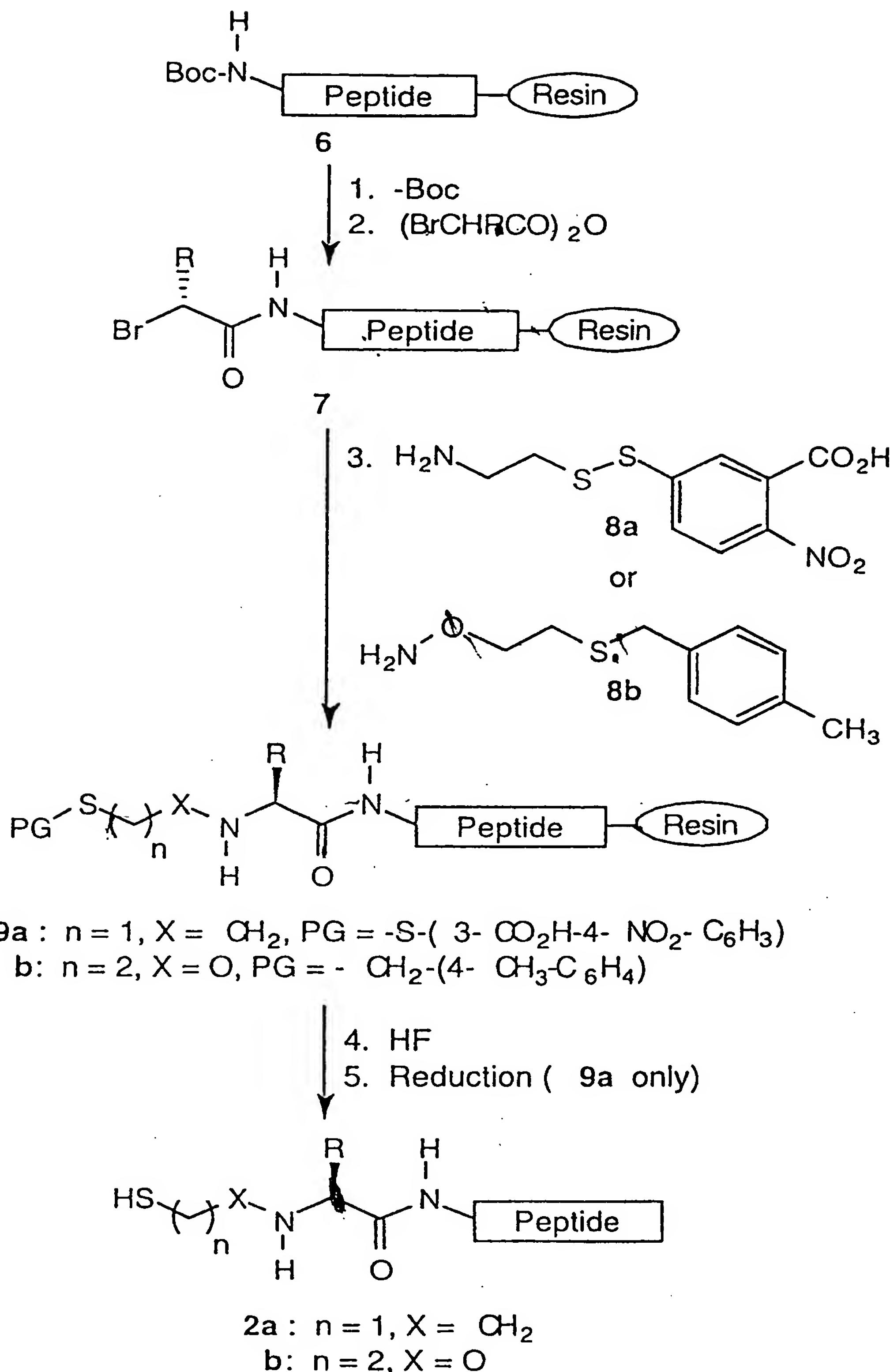
**FIG. 2**
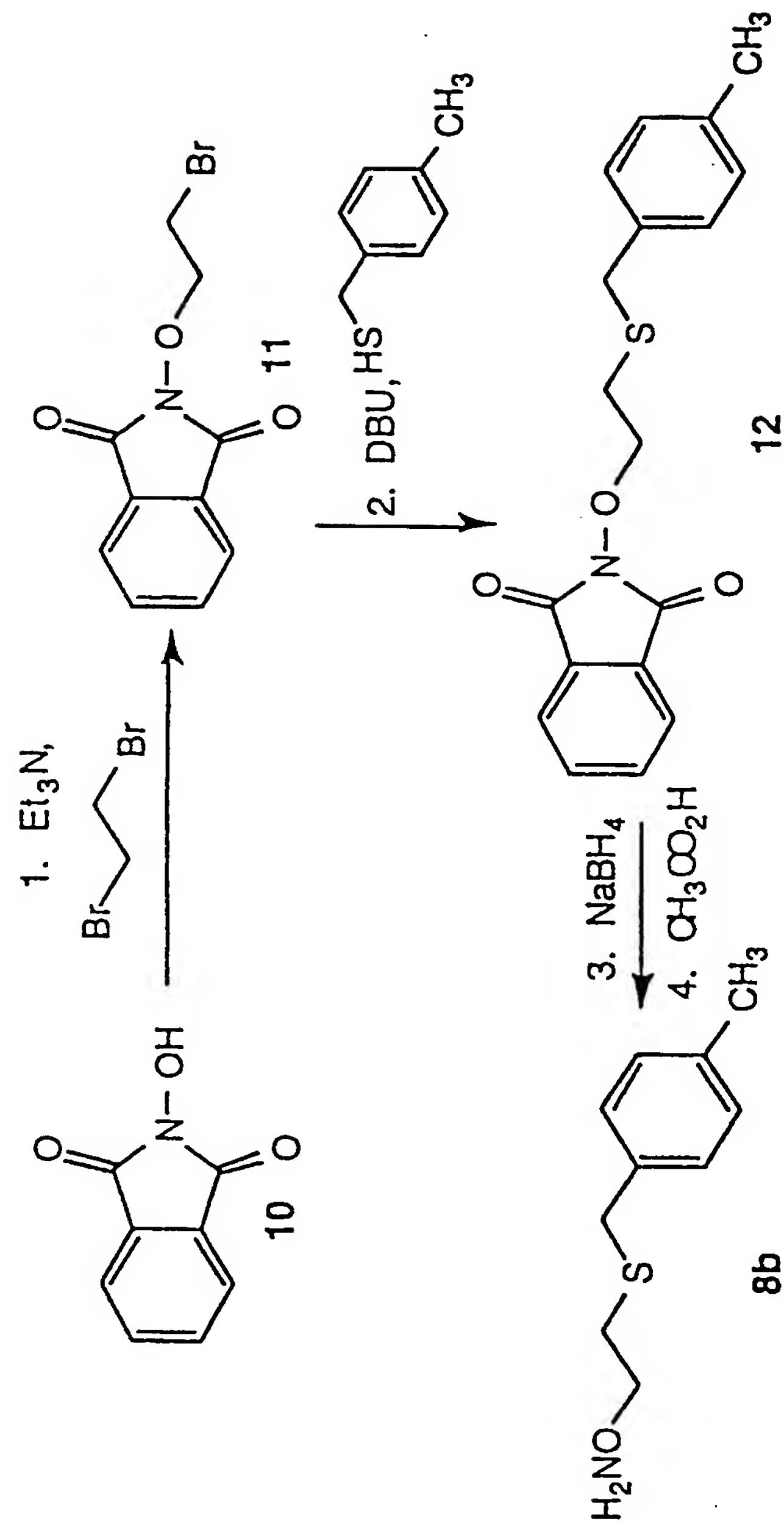
FIG. 3

| Example Ligation | Peptide 1[a] | Peptide 2[b] | pH | Temp. (°C) | Reaction time (hrs) | Approximate Yield[c] | |
|---|---|---|---|---|---|---|---|
| | | | | | | Unrearranged product (3) | Rearranged product (4) |
| 1 | LYRAG-SNB | Nα(etsh)-GAGPAGD | 7.0 | 25 | 4 | not detected | 90% |
| 2 | LYRAG-SNB | Nα(oetsh)-GRNTATIMMQRGNFR | 7.5 | 25 | 16 | not detected | 75% |
| 3 | LYRAF-Sbzl | Nα(oetsh)-GRNTATIMMQRGNFR | [7.5<br>4.5d | 37<br>37 | 11.5<br>10 | 30%<br>0% | (35%)]<br>64% |
| 4 | LYRAG-Sbzl | Nα(oetsh)-AARHTVHQRHLHG | [7.5<br>4.5d | 37<br>37 | 17.5<br>6.5 | 39%<br>20% | (52%)]<br>69% |
| 5 | LYRAF-Sbzl | Nα(oetsh)-AARHTVHQRHLHG | 7.5<br>4.5d | 37<br>37 | 19<br>22 | 58%<br>52% | not detected<br>not detected |

FIG. 4
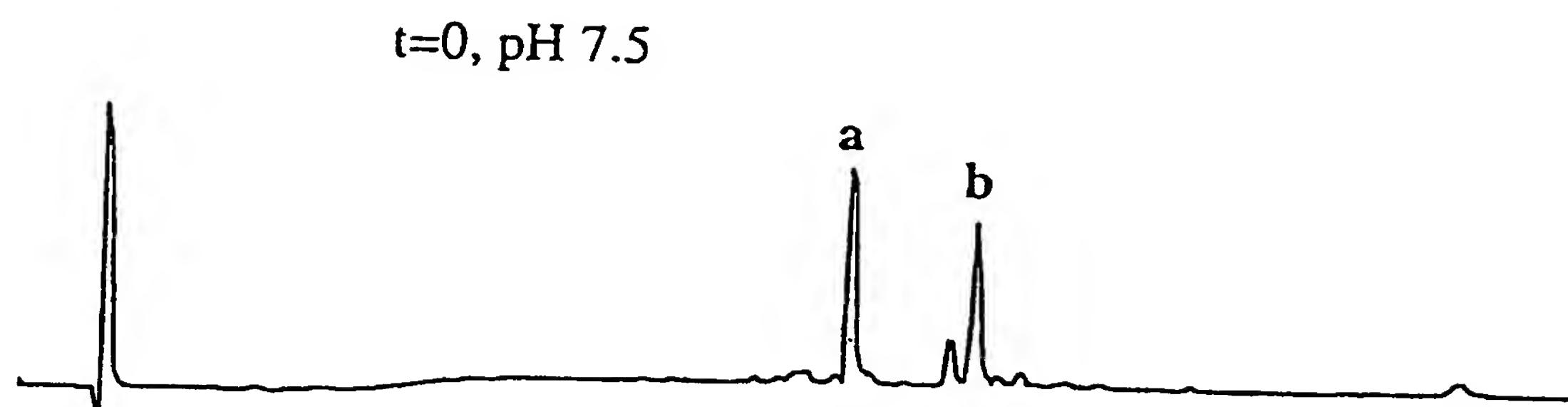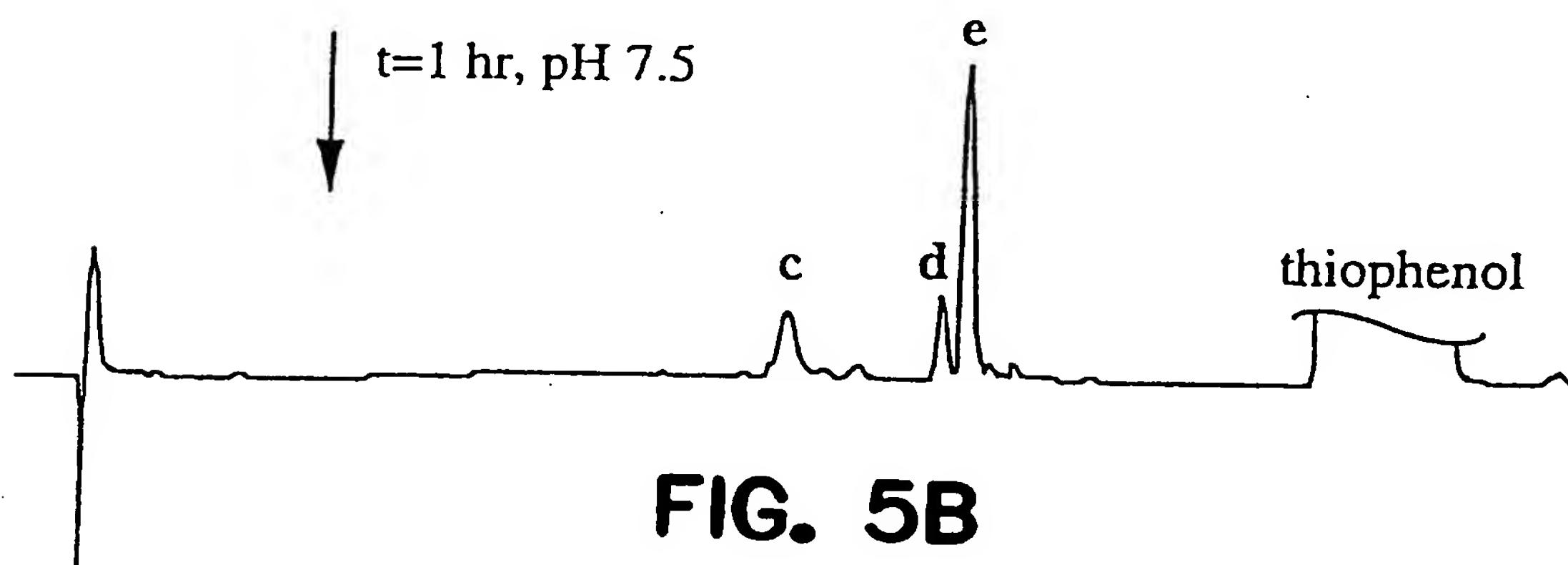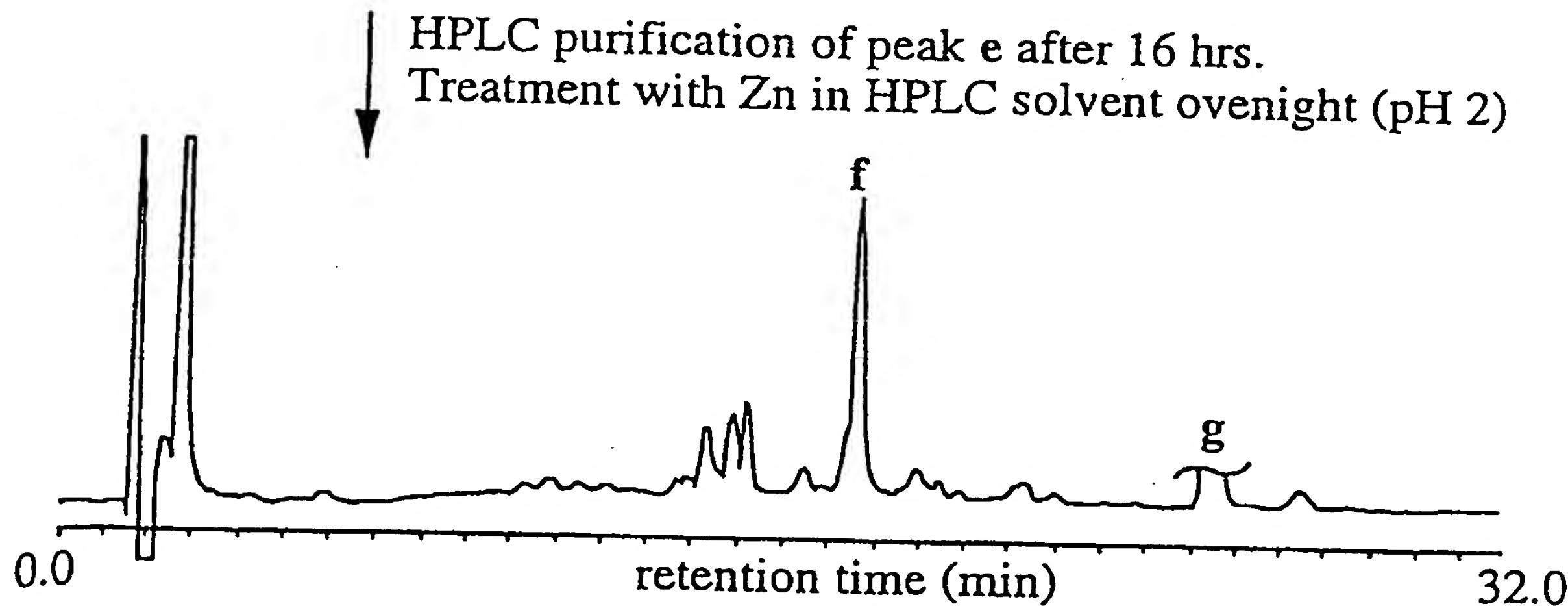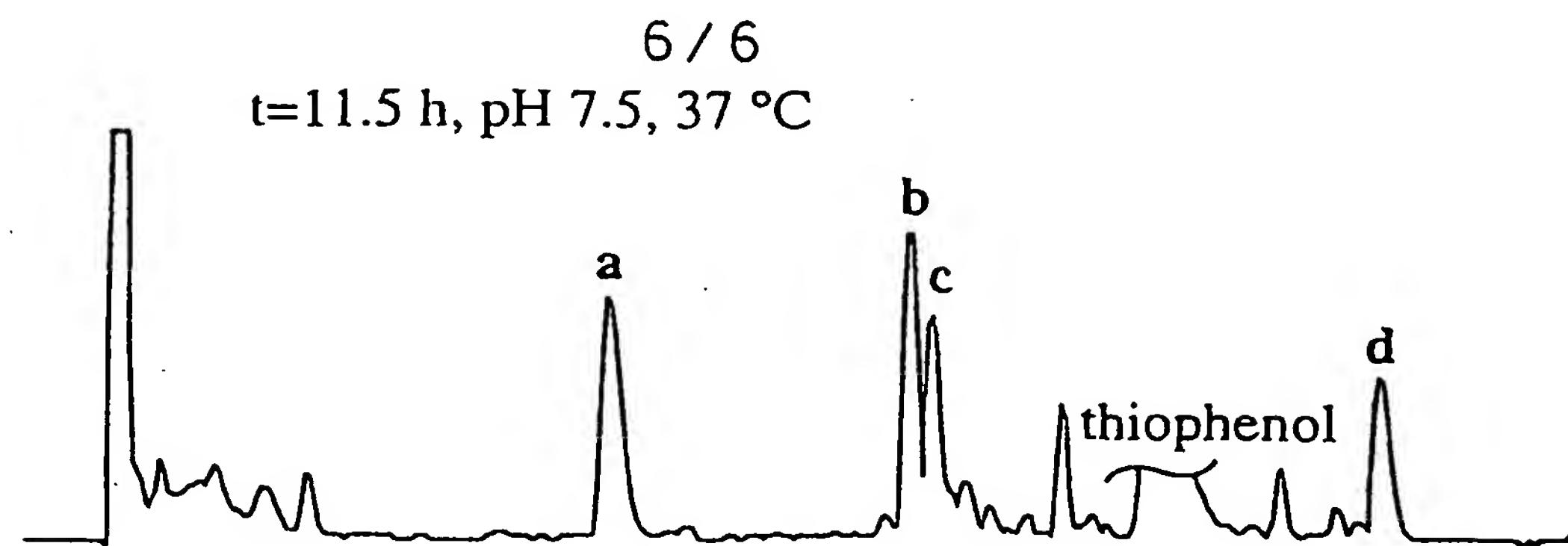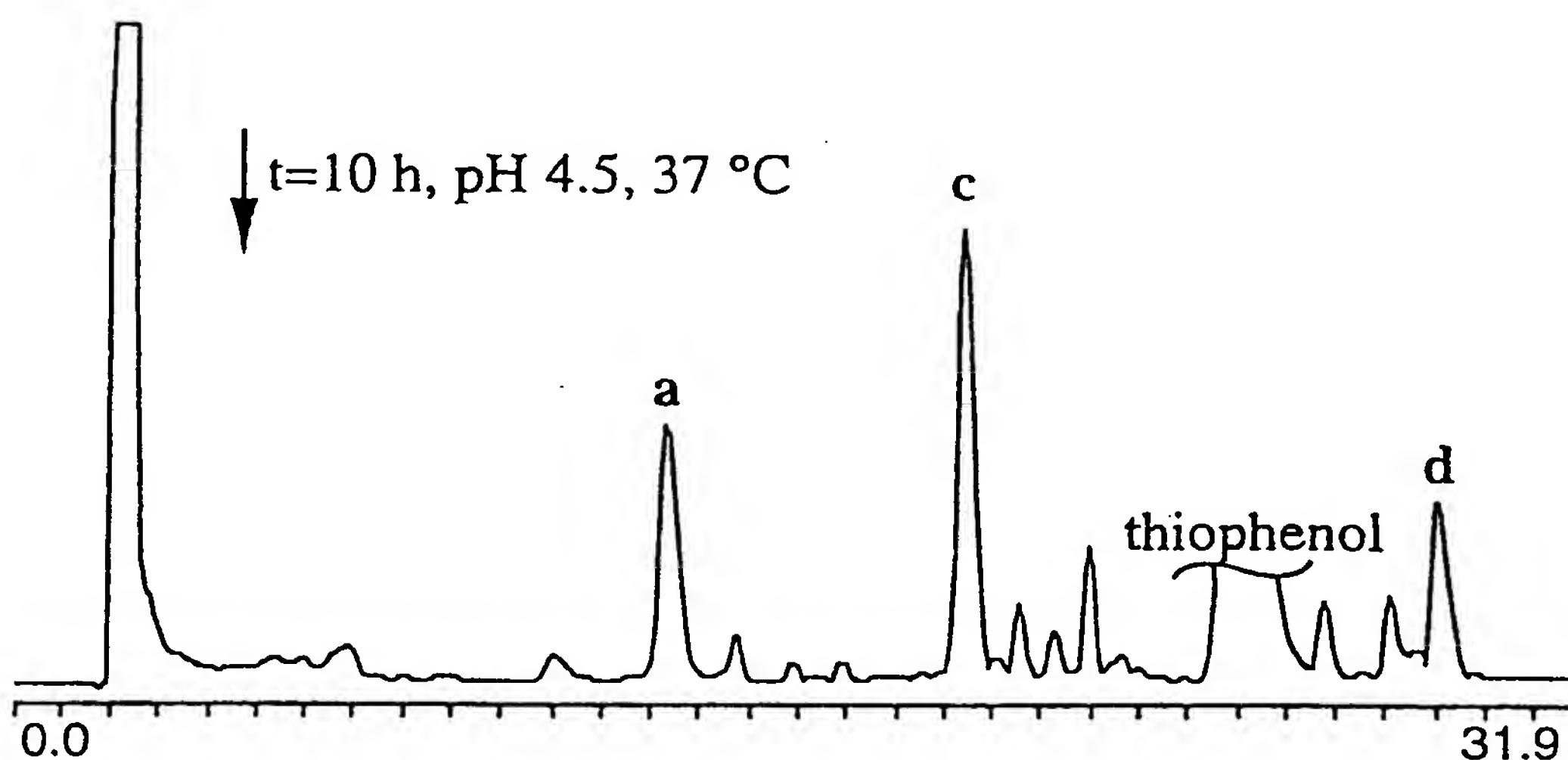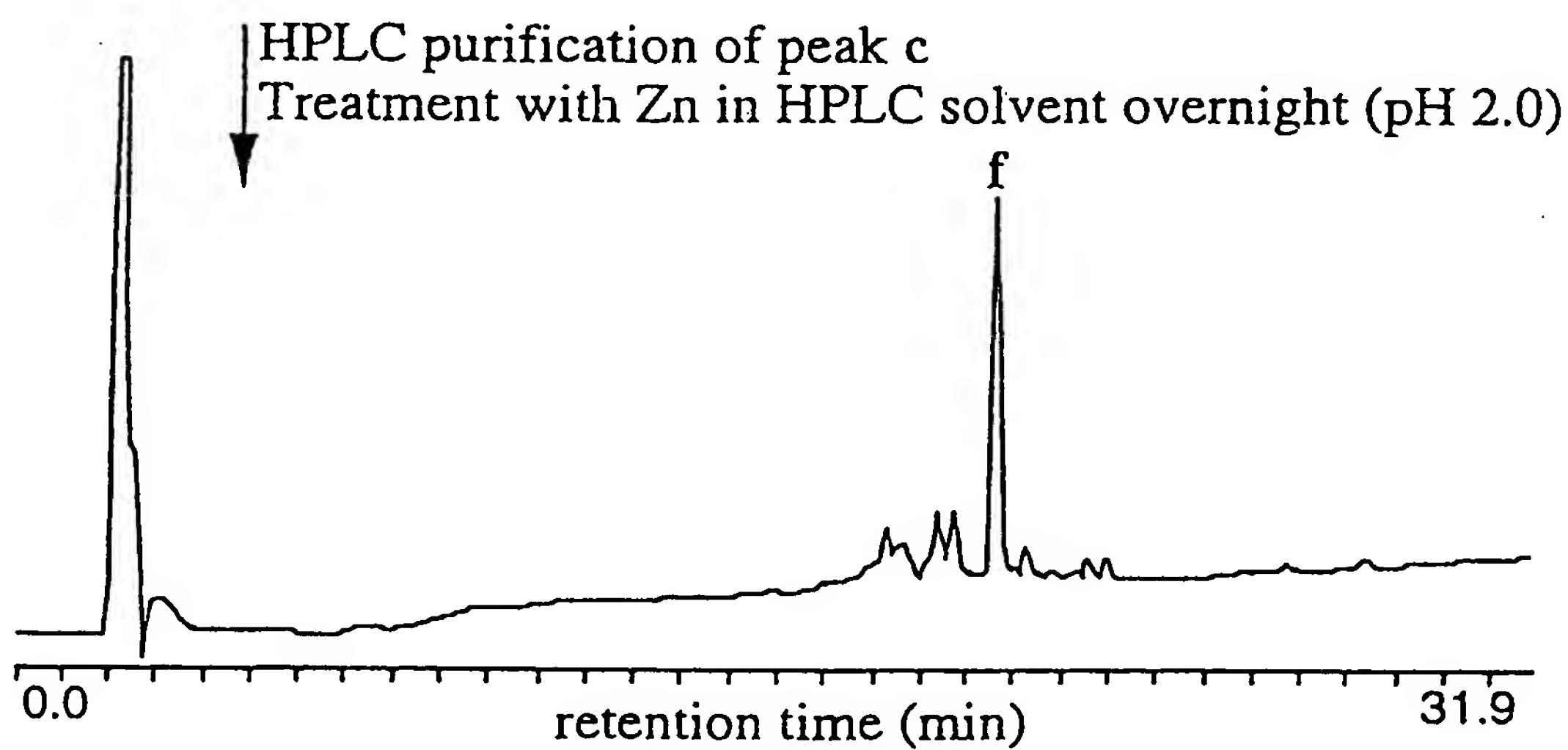
t=0, pH 7.5



**FIG. 5A**

t=1 hr, pH 7.5



**FIG. 5B**

HPLC purification of peak e after 16 hrs.
Treatment with Zn in HPLC solvent ovenight (pH 2)



0.0                    retention time (min)                    32.0

**FIG. 5C**

t=11.5 h, pH 7.5, 37 °C



**FIG. 6A**



**FIG. 6B**



retention time (min)

**FIG. 6C**

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(6)  :C12P 21/00, 21/06; C07K 1/00, 1/02, 1/04, 1/107, 2/00
US CL  :435/68.1; 530/300, 323, 333, 334, 338, 350, 402

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. :  435/68.1; 530/300, 323, 333, 334, 338, 350, 402

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | DAWSON et al.  Synthesis of proteins by native chemical ligation.  Science.  04 November 1994, Vol. 266, pages 776-779, especially pages 777-778 and figures 1-3. | 1-3, 7-8 |
| A, T | US 5,589,356 A (J.P. TAM) 31 December 1996 (31/12/96), see entire document. | 1-8 |
| A, T | US 5,625,030 A (WILLIAMS et al.) 29 April 1997 (29/04/97), see entire document. | 1-8 |

☐ Further documents are listed in the continuation of Box C.    ☐ See patent family annex.

| | | | |
|---|---|---|---|
| * | Special categories of cited documents: | "T" | later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention |
| "A" | document defining the general state of the art which is not considered to be of particular relevance | | |
| "E" | earlier document published on or after the international filing date | "X" | document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone |
| "L" | document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) | "Y" | document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art |
| "O" | document referring to an oral disclosure, use, exhibition or other means | | |
| "P" | document published prior to the international filing date but later than the priority date claimed | "&" | document member of the same patent family |

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 10 MAY 1997 | 0 2 JUN 1997 |

| Name and mailing address of the ISA/US | Authorized officer |
|---|---|
| Commissioner of Patents and Trademarks  Box PCT  Washington, D.C. 20231 | MICHAEL D. PAK |
| Facsimile No.    (703) 305-3230 | Telephone No.    (703) 308-0196 |

Form PCT/ISA/210 (second sheet)(July 1992)★

**B. FIELDS SEARCHED**
Electronic data bases consulted (Name of data base and where practicable terms used):

APS, BIOSCIENCE, CAPLUS, MEDLINE, SCISEARCH, EMBASE, WPIDS

search terms: peptide?(10a)synthesi?, peptide?(5a)ligat?, thioester?